

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 101718,311A
Source: FTW/6
Date Processed by STIC: 12/11/06

ENTERED



IFW16

RAW SEQUENCE LISTING

DATE: 12/11/2006

PATENT APPLICATION: US/10/718,311A

TIME: 12:31:00

Input Set : A:\BC1015USDIVcorrection_ST25.txt
Output Set: N:\CRF4\12112006\J718311A.raw

3 <110> APPLICANT: E.I. duPont de Nemours and Company, Inc.
4 Meyer, Knut
5 Viitanen, Paul
6 Van Dyk, Drew E.
8 <120> TITLE OF INVENTION: High Level Production of P-Hydroxybenzoic Acid in Green
Plants
10 <130> FILE REFERENCE: BC1015 US DIV
12 <140> CURRENT APPLICATION NUMBER: US 10/718,311A
13 <141> CURRENT FILING DATE: 2003-11-20
15 <160> NUMBER OF SEQ ID NOS: 18
17 <170> SOFTWARE: PatentIn version 3.4
19 <210> SEQ ID NO: 1
20 <211> LENGTH: 32
21 <212> TYPE: DNA
22 <213> ORGANISM: artificial sequence
24 <220> FEATURE:
25 <223> OTHER INFORMATION: Primer
27 <400> SEQUENCE: 1
28 ctactcattt catatgtcac accccgcgtt aa 32
31 <210> SEQ ID NO: 2
32 <211> LENGTH: 34
33 <212> TYPE: DNA
34 <213> ORGANISM: artificial sequence
36 <220> FEATURE:
37 <223> OTHER INFORMATION: Primer
39 <400> SEQUENCE: 2 34
40 catcttacta gatcttttagt acaacggta cggcc
43 <210> SEQ ID NO: 3
44 <211> LENGTH: 495
45 <212> TYPE: DNA
46 <213> ORGANISM: Escherichia coli
48 <400> SEQUENCE: 3
49 atgtcacacc ccgcgttaac gcaactgcgt gcgcgtgcgt attgtaaaga gatccctgcc 60
51 ctggatccgc aactgctcgta ctgggtgttg ctggaggatt ccatgacaaaa acgttttggaa 120
53 cagcagggaa aaacggtaag cgtgacgatg atccgcgaag ggtttgtcga gcagaatgaa 180
55 atccccgaag aactgccgct gctgccgaaa gagtctcggt actggttacg tgaaattttg 240
57 ttatgtgccg atggtaacc gtggcttgcc ggtcgtaaccg tcgttctgt gtcaacggtt 300
59 agcgggcccgg agctggcggtt acaaaaatttgg taaaacgc cgtaggacg ctatcttttc 360
61 acatcatcgat cattaacccg ggactttattt gagataggcc gtatgccgg gctgtggggg 420
63 cgacgttccc gcctgcgatt aagcggtaaa ccgctgttgc taacagaact gtttttaccg 480
65 gcgtcaccgt tgtac 495
68 <210> SEQ ID NO: 4
69 <211> LENGTH: 165
70 <212> TYPE: PRT

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/718,311A

DATE: 12/11/2006

TIME: 12:31:00

Input Set : A:\BC1015USDIVcorrection_ST25.txt
 Output Set: N:\CRF4\12112006\J718311A.raw

71 <213> ORGANISM: Escherichia coli
 73 <400> SEQUENCE: 4
 75 Met Ser His Pro Ala Leu Thr Gln Leu Arg Ala Leu Arg Tyr Cys Lys
 76 1 5 10 15
 79 Glu Ile Pro Ala Leu Asp Pro Gln Leu Leu Asp Trp Leu Leu Glu
 80 20 25 30
 83 Asp Ser Met Thr Lys Arg Phe Glu Gln Gln Gly Lys Thr Val Ser Val
 84 35 40 45
 87 Thr Met Ile Arg Glu Gly Phe Val Glu Gln Asn Glu Ile Pro Glu Glu
 88 50 55 60
 91 Leu Pro Leu Leu Pro Lys Glu Ser Arg Tyr Trp Leu Arg Glu Ile Leu
 92 65 70 75 80
 95 Leu Cys Ala Asp Gly Glu Pro Trp Leu Ala Gly Arg Thr Val Val Pro
 96 85 90 95
 99 Val Ser Thr Leu Ser Gly Pro Glu Leu Ala Leu Gln Lys Leu Gly Lys
 100 100 105 110
 103 Thr Pro Leu Gly Arg Tyr Leu Phe Thr Ser Ser Thr Leu Thr Arg Asp
 104 115 120 125
 107 Phe Ile Glu Ile Gly Arg Asp Ala Gly Leu Trp Gly Arg Arg Ser Arg
 108 130 135 140
 111 Leu Arg Leu Ser Gly Lys Pro Leu Leu Leu Thr Glu Leu Phe Leu Pro
 112 145 150 155 160
 115 Ala Ser Pro Leu Tyr
 116 165
 119 <210> SEQ ID NO: 5
 120 <211> LENGTH: 39
 121 <212> TYPE: DNA
 122 <213> ORGANISM: artificial sequence
 124 <220> FEATURE:
 125 <223> OTHER INFORMATION: Primer
 127 <400> SEQUENCE: 5
 128 ctactcaatt agatctccat ggcttcctct gtcatttct 39
 131 <210> SEQ ID NO: 6
 132 <211> LENGTH: 32
 133 <212> TYPE: DNA
 134 <213> ORGANISM: artificial sequence
 136 <220> FEATURE:
 137 <223> OTHER INFORMATION: Primer
 139 <400> SEQUENCE: 6
 140 cattttacat atatgccaca cctgcatgca gc 32
 143 <210> SEQ ID NO: 7
 144 <211> LENGTH: 684
 145 <212> TYPE: DNA
 146 <213> ORGANISM: artificial sequence
 148 <220> FEATURE:
 149 <223> OTHER INFORMATION: Chimeric gene encoding chloroplast-targeted CPL fusion protein
 151 <400> SEQUENCE: 7
 152 atggcttcct ctgtcatttc ttccagcagct gttgccacac gcagcaatgt tacacaagct 60
 154 agcatggttg caccttcac tggctctaaa tcttcagcca ctttccctgt tacaagaag 120

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/718,311A

DATE: 12/11/2006
TIME: 12:31:00

Input Set : A:\BC1015USDIVcorrection_ST25.txt
Output Set: N:\CRF4\12112006\J718311A.raw

156	caaaaaccttg	acatcaacttc	cattgctagc	aatggtggaa	gagtttagctg	catgcagggt	180										
158	tggcatatgt	cacaccccg	gttaacgc	ctgcgtgc	tgcgctattt	taaagagatc	240										
160	cctgccctgg	atccgcaact	gctcgactgg	ctgttgctgg	aggattccat	gacaaaacgt	300										
162	tttgaacagc	agggaaaaac	ggttaagcgt	acgatgatcc	gcgaagggtt	tgtcgagcag	360										
164	aatgaaatcc	ccgaagaact	gccgctgc	ccgaaagagt	ctcgtaactg	tttacgtgaa	420										
166	attttgttat	gtgccatgg	tgaaccgtgg	cttgcggtc	gtaccgtcgt	tcctgtgtca	480										
168	acgttaagcg	ggccggagct	gfcgttacaa	aaattgggt	aaacgccgtt	aggacgctat	540										
170	ctgttcacat	catcgacatt	aacccgggac	tttattgaga	taggcgtga	tgccggctg	600										
172	tgggggcac	gttccgcct	gcgattaagc	ggtaaacccgc	tgttgcta	acaactgttt	660										
174	ttaccggcgt	caccgtt	gtataa				684										
177	<210>	SEQ ID NO:	8														
178	<211>	LENGTH:	227														
179	<212>	TYPE:	PRT														
180	<213>	ORGANISM:	artificial sequence														
182	<220>	FEATURE:															
183	<223>	OTHER INFORMATION:	Synthetic chloroplast-targeted CPL fusion protein														
185	<400>	SEQUENCE:	8														
187	Met	Ala	Ser	Ser	Val	Ile	Ser	Ser	Ala	Ala	Val	Ala	Thr	Arg	Ser	Asn	
188	1				5				10					15			
191	Val	Thr	Gln	Ala	Ser	Met	Val	Ala	Pro	Phe	Thr	Gly	Leu	Lys	Ser	Ser	
192						20			25					30			
195	Ala	Thr	Phe	Pro	Val	Thr	Lys	Lys	Gln	Asn	Leu	Asp	Ile	Thr	Ser	Ile	
196						35			40					45			
199	Ala	Ser	Asn	Gly	Gly	Arg	Val	Ser	Cys	Met	Gln	Val	Trp	His	Met	Ser	
200						50			55					60			
203	His	Pro	Ala	Leu	Thr	Gln	Leu	Arg	Ala	Leu	Arg	Tyr	Cys	Lys	Glu	Ile	
204						65			70			75		80			
207	Pro	Ala	Leu	Asp	Pro	Gln	Leu	Leu	Asp	Trp	Leu	Leu	Glu	Asp	Ser		
208						85			90					95			
211	Met	Thr	Lys	Arg	Phe	Glu	Gln	Gly	Lys	Thr	Val	Ser	Val	Thr	Met		
212						100			105					110			
215	Ile	Arg	Glu	Gly	Phe	Val	Glu	Gln	Asn	Glu	Ile	Pro	Glu	Glu	Leu	Pro	
216						115			120					125			
219	Leu	Leu	Pro	Lys	Glu	Ser	Arg	Tyr	Trp	Leu	Arg	Glu	Ile	Leu	Leu	Cys	
220						130			135					140			
223	Ala	Asp	Gly	Glu	Pro	Trp	Leu	Ala	Gly	Arg	Thr	Val	Val	Pro	Val	Ser	
224						145			150			155		160			
227	Thr	Leu	Ser	Gly	Pro	Glu	Leu	Ala	Leu	Gln	Lys	Leu	Gly	Lys	Thr	Pro	
228						165			170					175			
231	Leu	Gly	Arg	Tyr	Leu	Phe	Thr	Ser	Ser	Thr	Leu	Thr	Arg	Asp	Phe	Ile	
232						180			185					190			
235	Glu	Ile	Gly	Arg	Asp	Ala	Gly	Leu	Trp	Gly	Arg	Arg	Ser	Arg	Leu	Arg	
236						195			200					205			
239	Leu	Ser	Gly	Lys	Pro	Leu	Leu	Leu	Thr	Glu	Leu	Phe	Leu	Pro	Ala	Ser	
240						210			215					220			
243	Pro	Leu	Tyr														
244	225																
247	<210>	SEQ ID NO:	9														
248	<211>	LENGTH:	34														

RAW SEQUENCE LISTING DATE: 12/11/2006
 PATENT APPLICATION: US/10/718,311A TIME: 12:31:00

Input Set : A:\BC1015USDIVcorrection_ST25.txt
 Output Set: N:\CRF4\12112006\J718311A.raw

```

249 <212> TYPE: DNA
250 <213> ORGANISM: artificial sequence
252 <220> FEATURE:
253 <223> OTHER INFORMATION: Primer
255 <400> SEQUENCE: 9
256 ctactcatt gaagactgca tgcagggttg gcat 34
259 <210> SEQ ID NO: 10
260 <211> LENGTH: 34
261 <212> TYPE: DNA
262 <213> ORGANISM: artificial sequence
264 <220> FEATURE:
265 <223> OTHER INFORMATION: Primer
267 <400> SEQUENCE: 10
268 catcttactg tcgacttttag tacaacggtg acgc 34
271 <210> SEQ ID NO: 11
272 <211> LENGTH: 37
273 <212> TYPE: DNA
274 <213> ORGANISM: artificial sequence
276 <220> FEATURE:
277 <223> OTHER INFORMATION: Primer
279 <400> SEQUENCE: 11
280 ctactcatt ggccagctct gtcatttctt cagcagc 37
283 <210> SEQ ID NO: 12
284 <211> LENGTH: 31
285 <212> TYPE: DNA
286 <213> ORGANISM: artificial sequence
288 <220> FEATURE:
289 <223> OTHER INFORMATION: Primer
291 <400> SEQUENCE: 12
292 catcttacta gatctttagt acaacggtga c 31
295 <210> SEQ ID NO: 13
296 <211> LENGTH: 33
297 <212> TYPE: DNA
298 <213> ORGANISM: artificial sequence
300 <220> FEATURE:
301 <223> OTHER INFORMATION: Primer
303 <400> SEQUENCE: 13
304 cccggggta cctaaagaag gagtgcgctcg aag 33
307 <210> SEQ ID NO: 14
308 <211> LENGTH: 46
309 <212> TYPE: DNA
310 <213> ORGANISM: artificial sequence
312 <220> FEATURE:
313 <223> OTHER INFORMATION: Primer
315 <400> SEQUENCE: 14
316 gatatcaagc tttcttagt cgacatcgat ctagtaacat agatga 46
319 <210> SEQ ID NO: 15
320 <211> LENGTH: 62
321 <212> TYPE: PRT

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/718,311A

DATE: 12/11/2006

TIME: 12:31:00

Input Set : A:\BC1015USDIVcorrection_ST25.txt

Output Set: N:\CRF4\12112006\J718311A.raw

322 <213> ORGANISM: artificial sequence
 324 <220> FEATURE:
 325 <223> OTHER INFORMATION: Synthetic chloroplast-targeting sequence
 327 <400> SEQUENCE: 15
 329 Met Ala Ser Ser Val Ile Ser Ser Ala Ala Val Ala Thr Arg Ser Asn
 330 1 5 10 15
 333 Val Thr Gln Ala Ser Met Val Ala Pro Phe Thr Gly Leu Lys Ser Ser
 334 20 25 30
 337 Ala Thr Phe Pro Val Thr Lys Lys Gln Asn Leu Asp Ile Thr Ser Ile
 338 35 40 45
 341 Ala Ser Asn Gly Gly Arg Val Ser Cys Met Gln Val Trp His
 342 50 55 60
 345 <210> SEQ ID NO: 16
 346 <211> LENGTH: 170
 347 <212> TYPE: PRT
 348 <213> ORGANISM: artificial sequence
 350 <220> FEATURE:
 351 <223> OTHER INFORMATION: Processed chloroplast-targeted CPL synthetic fusion protein
 353 <400> SEQUENCE: 16
 355 Met Gln Val Trp His Met Ser His Pro Ala Leu Thr Gln Leu Arg Ala
 356 1 5 10 15
 359 Leu Arg Tyr Cys Lys Glu Ile Pro Ala Leu Asp Pro Gln Leu Leu Asp
 360 20 25 30
 363 Trp Leu Leu Glu Asp Ser Met Thr Lys Arg Phe Glu Gln Gln Gly
 364 35 40 45
 367 Lys Thr Val Ser Val Thr Met Ile Arg Glu Gly Phe Val Glu Gln Asn
 368 50 55 60
 371 Glu Ile Pro Glu Glu Leu Pro Leu Leu Pro Lys Glu Ser Arg Tyr Trp
 372 65 70 75 80
 375 Leu Arg Glu Ile Leu Leu Cys Ala Asp Gly Glu Pro Trp Leu Ala Gly
 376 85 90 95
 379 Arg Thr Val Val Pro Val Ser Thr Leu Ser Gly Pro Glu Leu Ala Leu
 380 100 105 110
 383 Gln Lys Leu Gly Lys Thr Pro Leu Gly Arg Tyr Leu Phe Thr Ser Ser
 384 115 120 125
 387 Thr Leu Thr Arg Asp Phe Ile Glu Ile Gly Arg Asp Ala Gly Leu Trp
 388 130 135 140
 391 Gly Arg Arg Ser Arg Leu Arg Leu Ser Gly Lys Pro Leu Leu Leu Thr
 392 145 150 155 160
 395 Glu Leu Phe Leu Pro Ala Ser Pro Leu Tyr
 396 165 170
 399 <210> SEQ ID NO: 17
 400 <211> LENGTH: 180
 401 <212> TYPE: PRT
 402 <213> ORGANISM: Solanum lycopersicum
 404 <400> SEQUENCE: 17
 406 Met Ala Ser Ser Val Ile Ser Ser Ala Ala Val Ala Thr Arg Ser Asn
 407 1 5 10 15
 410 Val Thr Gln Ala Ser Met Val Ala Pro Phe Thr Gly Leu Lys Ser Ser

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/718,311A

DATE: 12/11/2006

TIME: 12:31:01

Input Set : A:\BC1015USDIVcorrection_ST25.txt

Output Set: N:\CRF4\12112006\J718311A.raw